Claims

We claim:

5   1.   In a communication system, a method for information dissemination, the method comprising the steps of:

providing a set of subscriptions, at least one of the set of subscriptions comprising a tree pattern, wherein the tree pattern comprises one or more interconnected nodes having a hierarchy and adapted to specify content and structure

10   of information; and

using the set of subscriptions to select information for dissemination to one or more users.

2.   The method of claim 1, wherein the at least one subscription describes

15   information the one or more users are interested in receiving.

3.   The method of claim 1, further comprising the step of determining an aggregation from the set of subscriptions, the aggregation comprising a set of aggregate patterns, wherein the set of aggregate patterns is smaller than the set of

20   subscriptions, and wherein the step of using the set of subscriptions to select information for dissemination further comprises using the set of aggregate patterns to select the information for dissemination to the one or more users.

4.   The method of claim 1, wherein the information comprises one or

25   more documents defined using extensible markup language (XML).

5.   The method of claim 3, wherein at least one of the aggregate patterns and the tree pattern each is defined using extensible markup language (XML).

30   6.   The method of claim 3, wherein each aggregate pattern and each subscription comprises a tree pattern having one or more interconnected nodes having a hierarchy, and wherein the set of aggregate patterns is smaller than the set of

subscriptions in that a number of aggregate patterns in the set of aggregate patterns is smaller than a number of tree patterns in the set of subscriptions and that a number of nodes in the set of aggregate patterns is smaller than a number of nodes in the set of subscriptions.

5

7.   The method of claim 3, wherein the step of determining an aggregation further comprises the step of determining the aggregation from the set of subscriptions by using at least a space constraint.

10   8.   The method of claim 7, wherein the space constraint comprises a predetermined number of bytes.

9.   The method of claim 3, wherein the set of subscriptions comprises a plurality of tree patterns, each of the tree patterns comprising one or more

15   interconnected nodes having a hierarchy and adapted to specify content and structure of information, and wherein the step of determining an aggregation further comprises the step of determining a least upper bound pattern for two of the plurality of tree patterns in the set of subscriptions, the least upper bound pattern chosen as an aggregate pattern.

20

10.   The method of claim 9, wherein the two tree patterns are a first tree pattern and a second tree pattern, and wherein the step of determining a least upper bound pattern further comprises the steps of:

        if the first tree pattern is contained in the second tree pattern, setting

25   the least upper bound pattern to be the first tree pattern;

        if the second tree pattern is contained in the first tree pattern, setting the least upper bound pattern to be the second tree pattern;

        traversing the first and second tree patterns and computing a tightest container pattern by:

30   computing a position-preserving tightest container pattern by finding common sub-patterns;

computing an off-position tightest container pattern by finding common sub-patterns; and

constructing the tightest container pattern by taking a union of the position-preserving tightest container pattern and the off-position tightest container pattern,

wherein the tightest container pattern is used as the least upper bound pattern.

11.   The method of claim 9, wherein the step of determining a least upper bound pattern for two of the plurality of tree patterns further comprises the steps of determining a tightest container pattern for the two tree patterns and minimizing the tightest container pattern to create a minimal pattern, wherein the minimal pattern is used as the least upper bound pattern.

12.   The method of claim 3, wherein the set of subscriptions comprises a plurality of tree patterns, wherein each tree pattern in the set of subscriptions comprises one or more interconnected nodes having a hierarchy and adapted to specify content and structure of information, and wherein the step of determining an aggregation further comprises the steps of:

designating a candidate set of tree patterns to be the plurality of tree patterns;

performing the following steps:

identifying a set of candidate aggregate patterns from the plurality of tree patterns and similar tree patterns determined from the candidate set of tree patterns;

pruning each candidate aggregate pattern by deleting or merging nodes;

selecting a chosen tree pattern from the candidate aggregate patterns having a predetermined marginal gain; and

replacing all tree patterns, in the candidate set of tree patterns, that are contained in the chosen tree pattern by the chosen tree pattern.

13.   The method of claim 12, wherein the marginal gain is determined by a benefit value of a tree pattern.

5   14.   The method of claim 13, wherein the candidate set of tree patterns occupies a space and wherein the benefit value is determined from a ratio of savings in the space for a corresponding tree pattern to a loss in selectivity for the corresponding tree pattern.

10   15.   The method of claim 14, wherein the selectivity is determined by sampling matching of information to candidate patterns.

16.   In a communication system, an apparatus for providing information dissemination, the apparatus comprising:

15   a memory; and

             at least one processor, coupled to the memory;

             the apparatus operative:

             to provide a set of subscriptions, at least one of the set of subscriptions comprising a tree pattern, wherein the tree pattern comprises one or more

20   interconnected nodes having a hierarchy and adapted to specify content and structure of information; and

             to use the set of subscriptions to select information for dissemination to one or more users.

25   17.   An article of manufacture for providing information dissemination, the article of manufacture comprising:

             a machine readable medium containing one or more programs which when executed implement the steps of:

             · providing a set of subscriptions, at least one of the set of subscriptions

30   comprising a tree pattern, wherein the tree pattern comprises one or more interconnected nodes having a hierarchy and adapted to specify content and structure of information; and

using the set of subscriptions to select information for dissemination to one or more users.